

WKB calculation of an epidemic outbreak distribution

Andrew J Black¹ and Alan J McKane²

¹ School of Mathematical Sciences, The University of Adelaide, Adelaide, SA 5005, Australia

² Theoretical Physics Division, School of Physics and Astronomy, University of Manchester, Manchester M13 9PL, UK

E-mail: `andrew.black@adelaide.edu.au`

Abstract. We calculate both the exponential and pre-factor contributions in a WKB approximation of the master equation for a stochastic SIR model with highly oscillatory dynamics. Fixing the basic parameters of the model we investigate how the outbreak distribution changes with the population size. We show that this distribution rapidly becomes highly non-Gaussian, acquiring large tails indicating the presence of rare, but large outbreaks, as the population is made smaller. The analytic results are found to be in excellent agreement with simulations until the systems become so small that the dynamics are dominated by fade-out of the disease.

Keywords: Markov process, master equation, childhood diseases, stochastic amplification, WKB approximation.

Submitted to: *J. Stat. Mech.*

1. Introduction

The dynamics of recurrent epidemics are an active topic of research in a wide range of disciplines: from ecology and epidemiology to applied mathematics and theoretical physics [1–3]. Much of the reason for this can be attributed to the large amount of historical data available [4,5], and the rich variety of highly oscillatory patterns that are observed in both the data and relatively simple stochastic models [6,7].

The systematic formulation of individually-based models of epidemics starting from a master equation (continuous-time Markov chain) has permitted a parallel analytical and numerical study to be carried out, which has allowed a rather complete understanding of these stochastic models to be achieved. One of the main features of these models is that demographic stochasticity tends to excite the natural dynamics of the system leading to large scale coherent oscillations, termed stochastic amplification [8]. It has been shown how the frequency and amplitude of these oscillations depend on the parameters of the system and can provide a parsimonious explanation for the range of frequencies observed in real data [3,9–11]. In all of these studies though, the population is assumed to be very large so low numbers of infectives, or fade-out of the disease, is unlikely.

For childhood diseases especially, where the infectious period is orders of magnitude smaller than the lifetime [12], the size of the epidemic outbreaks can be so large that the number of infectives can become small and fade-out of the disease is very probable, even in medium-size populations [4]. Because of this, some form of migration is commonly included in these models to re-introduce the disease after fade-out. This is a biologically realistic feature which is not needed in deterministic models where fade-out cannot happen. The presence of this fade-out boundary in the stochastic model means there is a natural asymmetry in the state space, which can have a considerable effect on the dynamics of the fluctuations in smaller-to-medium size systems. Figure 1 illustrates the dynamics of a stochastic epidemic model at three different population sizes. As the population is decreased the probability of a large outbreak, relative to the mean infection level, increases. These smaller systems are important as they are representative of the majority of towns and cities for which extensive case-report data exists [13], but have so far received little theoretical attention.

In this paper we calculate analytically—via a Wentzel-Kramers-Brillouin (WKB) approximation of the master equation—the outbreak distribution (marginal infective probability distribution) of a susceptible-infected-recovered (SIR) model. We show how the asymmetry in the system affects this distribution and in particular how it changes with population size. In the limit $N \rightarrow \infty$, the distribution tends to a Gaussian, as is expected and used in the system-size expansion of the master equation [14]. As the population is made smaller the distribution becomes skewed, rapidly acquiring fat-tails indicating an increased probability of rare, but large, outbreaks. All results are compared with stochastic simulations and, within the approximation’s range of validity, the agreement is excellent. Although we have used the term analytic to refer to the

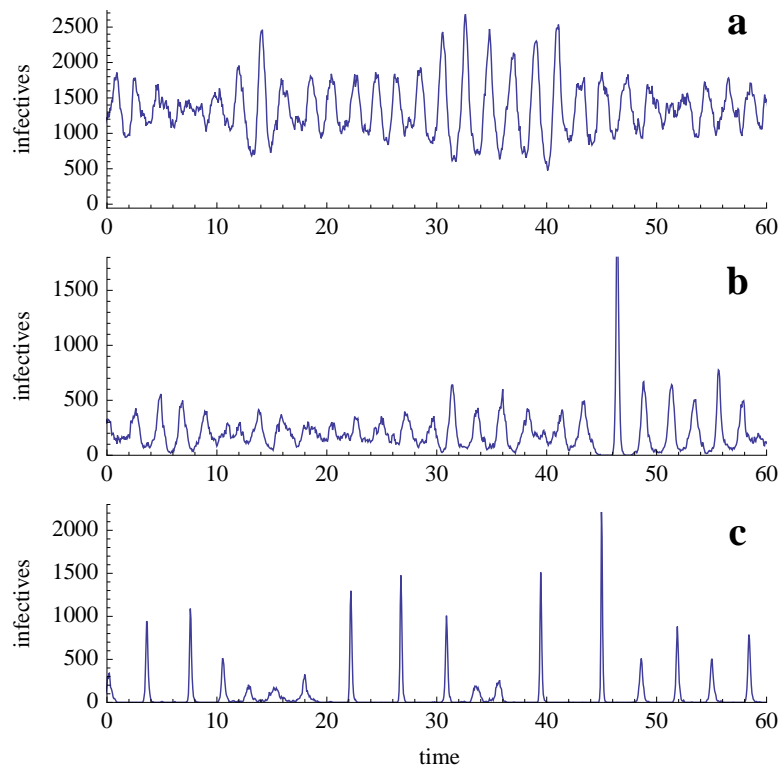


Figure 1. Infective time series illustrating how the dynamics of the model change as the total population, N , is decreased. (a) $N = 2 \times 10^6$, the number of infectives oscillates about the mean, but fade-out does not happen. (b) $N = 3 \times 10^5$, fade-out does occur but rarely. (c) $N = 10^5$, here outbreaks are mainly triggered by immigration events and rapidly lead to fade-out. The model and parameters which generate these are described in Sections 2 and 3.

method which we use, it should be understood by this that we use analytic techniques to obtain equations or expressions for functions making up probability distributions, but numerical techniques have to be used to solve these to obtain the final results.

The WKB approximation has a long history in the theory of stochastic processes, and was the basis of Kramer’s calculation of escape of a particle over a potential barrier due to thermal noise [15]. The method was originally used on Fokker-Planck equations with continuous variables \mathbf{x} and consisted of making an ansatz for the probability of the form $P(\mathbf{x}) \sim \exp[-NS(\mathbf{x})]$, where N was a large parameter such as the volume of the system or the inverse diffusion constant. The function $S(\mathbf{x})$ turns out to obey a Hamilton-Jacobi equation [16], and so leads to a definition of a Hamiltonian dynamics which describes the stochastic dynamics of rare events, such as large outbreaks.

For the situation discussed in this paper we require a variant of the method applied to the master equation, where the variables are discrete rather than continuous [17–19]. However the method is similar and the problem of describing strongly stochastic effects near the fade-out boundary is transformed into one of classical mechanics and of finding zero energy trajectories of a Hamiltonian. It should be stressed that this deterministic dynamics is not simply the $N \rightarrow \infty$ dynamics of the original stochastic model, but an

auxiliary dynamics which is able to give a very good description of processes that cannot be described by the conventional deterministic dynamics found by taking $N \rightarrow \infty$.

The method allows higher-order corrections to the leading $\exp[-NS(\mathbf{x})]$ to be calculated, the next-order correction being a prefactor $K(\mathbf{x})$ multiplying the exponential [20,21]. There have been a number of applications of this method in the last few years, for example to calculate fixation and extinction times from quasi-stationary to absorbing states [22–25], or transition times between metastable states [19, 26, 27]. The WKB method is especially suitable for these type of problems as the aim is to only find one special trajectory. The work in this paper differs from these previous studies in that we are primarily concerned in studying the dynamics of the model, and hence in calculating the full stationary distribution of the master equation, not fixation times (there is no fixation in our model since immigration is included). Another distinguishing aspect of our work is that we also calculate the pre-factor term for the multi-dimensional model, which as will be shown, is crucial for accurate calculation of the probability density. This has not so far been computed for a model of this type.

The rest of this paper is organised as follows: in Section 2 we introduce the theory behind the WKB approximation of the master equation and in Section 3 describe the numerical procedure that we adopt to apply it to calculate approximate distributions. In Section 4 we apply these methods to the SIR model with immigration. We end with a discussion of our results and outline a number of open questions.

2. General formalism

In this section we will discuss the WKB approximation to the master equation and its application to the SIR model. Although much of the formalism we require appears in various places in the literature it is somewhat scattered, so we first give a coherent summary based on some of the clearer discussions that have been published [19–21].

The starting point is the master equation for the time-evolution of $P_{\mathbf{n}}(t)$, which is the probability of finding the system in state \mathbf{n} at time t :

$$\frac{dP_{\mathbf{n}}(t)}{dt} = \sum_{\mathbf{r}} [T_{\mathbf{r}}(\mathbf{n} - \mathbf{r})P_{\mathbf{n}-\mathbf{r}}(t) - T_{\mathbf{r}}(\mathbf{n})P_{\mathbf{n}}(t)] . \quad (1)$$

Here \mathbf{r} labels the transitions by giving the size of the jump. Both \mathbf{n} and \mathbf{r} are vectors of integers; for the SIR model they are two-dimensional with $\mathbf{n} = (n, m)$, where n and m are the number of susceptible and infected individuals respectively. Setting the left-hand side of Equation (1) to zero we find the time-independent equation specifying the stationary distribution,

$$\sum_{\mathbf{r}} [T_{\mathbf{r}}(\mathbf{n} - \mathbf{r})P_{\mathbf{n}-\mathbf{r}} - T_{\mathbf{r}}(\mathbf{n})P_{\mathbf{n}}] = 0 . \quad (2)$$

We now expand the transition rates in terms of the system size, $N \gg 1$. In our case this is the number of individuals in the system. Specifically we write

$$T_{\mathbf{r}}(\mathbf{n}) = NW_{\mathbf{r}}(\mathbf{n}) + U_{\mathbf{r}}(\mathbf{n}) + \mathcal{O}(1/N) , \quad (3)$$

where $W_{\mathbf{r}}(\mathbf{n})$ and $U_{\mathbf{r}}(\mathbf{n})$ are of order unity. For the SIR Model, the transition rates are such that $U_{\mathbf{r}}(\mathbf{n})$ and the higher-order corrections are zero, and so $T_{\mathbf{r}} = NW_{\mathbf{r}}$.

Finally we introduce the fraction of individuals who fall into the different classes: $\mathbf{x} = \mathbf{n}/N$; for the SIR model we will write $\mathbf{x} = (x, y)$ where $x = n/N$ and $y = m/N$. So letting $W_{\mathbf{r}}(\mathbf{n}) = W_{\mathbf{r}}(N\mathbf{x}) = w_{\mathbf{r}}(\mathbf{x})$ and also $P_{\mathbf{n}} \equiv P_{N\mathbf{x}} = \pi(\mathbf{x})$, the master equation (2) now reads

$$\sum_{\mathbf{r}} \left[w_{\mathbf{r}}(\mathbf{x} - \frac{\mathbf{r}}{N}) \pi(\mathbf{x} - \frac{\mathbf{r}}{N}) - w_{\mathbf{r}}(\mathbf{x}) \pi(\mathbf{x}) \right] = 0. \quad (4)$$

2.1. The WKB approximation

We can now apply the WKB approximation to $\pi(\mathbf{x})$ by expressing it as

$$\pi(\mathbf{x}) = K(\mathbf{x}) \exp(-NS(\mathbf{x})) \left[1 + \mathcal{O}\left(\frac{1}{N}\right) \right], \quad (5)$$

where both $S(\mathbf{x})$ and $K(\mathbf{x})$ are assumed to be of order unity and $N \gg 1$. Next we expand $S(\mathbf{x} - (\mathbf{r}/N))$ to second order:

$$S(\mathbf{x} - \frac{\mathbf{r}}{N}) = S(\mathbf{x}) - \frac{1}{N} \mathbf{r} \cdot \nabla_{\mathbf{x}} S(\mathbf{x}) + \frac{1}{2N^2} (\mathbf{r} \cdot \nabla_{\mathbf{x}})^2 S(\mathbf{x}) + \mathcal{O}\left(\frac{1}{N^3}\right), \quad (6)$$

and $K(\mathbf{x})$ to first order: $K(\mathbf{x} - \mathbf{r}/N) = K(\mathbf{x}) - \mathbf{r} \cdot \nabla_{\mathbf{x}} K(\mathbf{x})/N + \mathcal{O}(1/N^2)$. Substituting into the master equation and comparing terms, at leading order we find a Hamilton-Jacobi equation, corresponding to the Hamiltonian

$$H(\mathbf{x}, \mathbf{p}) = \sum_{\mathbf{r}} w_{\mathbf{r}}(\mathbf{x}) [\exp(\mathbf{r} \cdot \mathbf{p}) - 1] = 0. \quad (7)$$

Here $\mathbf{p} = \nabla_{\mathbf{x}} S(\mathbf{x})$. The prefactor $K(\mathbf{x})$ is determined by the next-order contributions in N^{-1} and is discussed in Section 2.3. It should not be too surprising that a Hamilton-Jacobi equation is found using the form (5): Markovian stochastic processes are equivalent to quantum mechanics in imaginary time [28] and we are interested in the limit where N^{-1} (\hbar) goes to zero.

Hamilton's equations are found from Eq. (7) to be:

$$\begin{aligned} \dot{\mathbf{x}} &= \nabla_{\mathbf{p}} H = \sum_{\mathbf{r}} \mathbf{r} w_{\mathbf{r}}(\mathbf{x}) \exp(\mathbf{r} \cdot \mathbf{p}), \\ \dot{\mathbf{p}} &= -\nabla_{\mathbf{x}} H = -\sum_{\mathbf{r}} [\exp(\mathbf{r} \cdot \mathbf{p}) - 1] \nabla_{\mathbf{x}} w_{\mathbf{r}}(\mathbf{x}), \end{aligned} \quad (8)$$

where the dot denotes differentiation with respect to time. From the solution of these equations we will find trajectories, called the fluctuational trajectories, $\mathbf{x}_f(t)$ and the “momenta” along these trajectories $\mathbf{p}_f(t)$. For the zero-energy solutions that we are interested in, the action along a given fluctuational trajectory is given by

$$S_f = \int_{t_0}^t \mathbf{p}_f \cdot \dot{\mathbf{x}}_f dt'. \quad (9)$$

From Eq. (8) we see that $\mathbf{p}_f = 0$ is always one solution, with

$$\dot{\mathbf{x}} = \sum_{\mathbf{r}} \mathbf{r} w_{\mathbf{r}}(\mathbf{x}). \quad (10)$$

This is called the relaxation trajectory for the following reason. If we multiply the original master equation (1) by \mathbf{n} , sum over all \mathbf{n} , and then shift the first sum by \mathbf{r} , we find that $d\langle\mathbf{n}\rangle/dt = \sum_{\mathbf{r}} \mathbf{r} T_{\mathbf{r}}(\langle\mathbf{n}\rangle)$, in the limit $N \rightarrow \infty$. Introducing \mathbf{x} and $w_{\mathbf{r}}(\mathbf{x})$ as above, gives exactly Eq. (10). Therefore Eq. (10) describes the actual deterministic dynamics of the system which would occur in the limit $N \rightarrow \infty$. This is not the trajectory of primary interest to us in this paper; instead our focus will be on solutions of Eq. (8) with $\mathbf{p}_f \neq 0$, which correspond to non-allowed trajectories in the $N \rightarrow \infty$ limit. Before we discuss these, we will apply this formalism to the SIR model.

2.2. The SIR model with immigration

The SIR model consists of three classes of individuals, susceptibles denoted by S , infected denoted by I and recovered denoted by R [12]. We assume that births and deaths are coupled at the individual level, so that when an individual dies another (susceptible) individual is born [8]. This means that the number of individuals, N , does not change with time, and so the number of recovered individuals is not an independent variable: $n_R = N - n - m$. Immigration is included with a commuter formalism where susceptibles are assumed to be in contact with a constant pool of infectives outside the main community [7, 8].

The transitions in the model are then of the following type:

- (i) Infection: $S + I \xrightarrow{\beta} I + I$ and $S \xrightarrow{\eta} I$.

$$T(n-1, m+1|n, m) = \left(\beta \frac{n}{N} m + \eta n\right).$$

- (ii) Recovery: $I \xrightarrow{\gamma} R$.

$$T(n, m-1|n, m) = \gamma m.$$

- (iii) Death of an infected individual: $I \xrightarrow{\mu} S$.

$$T(n+1, m-1|n, m) = \mu m.$$

- (iv) Death of a recovered individual: $R \xrightarrow{\mu} S$.

$$T(n+1, m|n, m) = \mu(N - n - m).$$

Expressing this model in terms of the quantities introduced at the start of this section, we find that, as already mentioned, all corrections to $NW_{\mathbf{r}}(\mathbf{n})$ in Eq. (3) are zero, and that $w_{\mathbf{r}}(\mathbf{x})$ is given by

- (i) Infection: $w_{\mathbf{r}}(\mathbf{x}) = \beta xy + \eta x$, $\mathbf{r} = (-1, +1)$

- (ii) Recovery: $w_{\mathbf{r}}(\mathbf{x}) = \gamma y$, $\mathbf{r} = (0, -1)$

- (iii) Death of an infected individual: $w_{\mathbf{r}}(\mathbf{x}) = \mu y$, $\mathbf{r} = (+1, -1)$

- (iv) Death of a recovered individual: $w_{\mathbf{r}}(\mathbf{x}) = \mu(1 - x - y)$, $\mathbf{r} = (+1, 0)$

Substituting these rates into (7) we find the Hamiltonian

$$H(\mathbf{x}, \mathbf{p}) = (\beta xy + \eta x)[e^{-p_x + p_y} - 1] + \gamma y[e^{-p_y} - 1] + \mu y[e^{-p_y + p_x} - 1] + \mu(1 - x - y)[e^{p_x} - 1]. \quad (11)$$

Related Hamiltonians have been used by previous authors [22, 24]. The relaxation trajectory can be found by setting $\mathbf{p} = 0$ in the Hamilton equations found from Equation (11), or by simply substituting the $w_{\mathbf{r}}(\mathbf{x})$ for the SIR model into Equation (10). Either way this gives the deterministic SIR equations:

$$\begin{aligned} \dot{x} &= -\beta xy - \eta x + \mu(1 - x), \\ \dot{y} &= \beta xy + \eta x - (\gamma + \mu)y. \end{aligned} \quad (12)$$

From Equation (12) we see that the fixed points (denoted by an asterisk) satisfy $y^* = \mu(1 - x^*)/(\gamma + \mu)$. Eliminating y^* shows that x^* may be found by solving

$$(x^*)^2 - \left[1 + \frac{\eta(\gamma + \mu)}{\mu\beta} + \frac{(\gamma + \mu)}{\beta}\right]x^* + \frac{(\gamma + \mu)}{\beta} = 0. \quad (13)$$

This equation has two fixed points: a ‘trivial’ one which vanishes if $\eta = 0$ and a non-trivial (endemic) one. This is the one which will be of interest to us in the rest of the paper. The focus will be on the fluctuational trajectories which after a perturbation to this fixed point at $t = 0$, rapidly move away and end near the extinction boundary [19, 23].

2.3. Calculation of the Prefactor $K(\mathbf{x})$

An equation for the prefactor $K(\mathbf{x})$ is found from examining the second-order $\mathcal{O}(N^{-1})$ terms in the expansion of the master equation. From this we find the equation

$$\sum_{\mathbf{r}} w_{\mathbf{r}} \exp(\mathbf{r} \cdot \mathbf{p}) \left[\frac{\mathbf{r} \cdot \nabla_{\mathbf{x}} K}{K} + \frac{1}{2} (\mathbf{r} \cdot \nabla_{\mathbf{x}})^2 S + \frac{\mathbf{r} \cdot \nabla_{\mathbf{x}} w_{\mathbf{r}}}{w_{\mathbf{r}}} \right] = 0. \quad (14)$$

If $U_r \neq 0$ in the expansion of the transition rates (3), then there are extra terms in this equation [27]. Using Hamilton’s equations (8), and their derivatives, we may write Equation (14) in the alternative form

$$\frac{1}{K} \sum_i \frac{\partial H}{\partial p_i} \frac{\partial K}{\partial x_i} + \frac{1}{2} \sum_{i,j} \frac{\partial^2 S}{\partial x_i \partial x_j} \frac{\partial^2 H}{\partial p_i \partial p_j} + \sum_i \frac{\partial^2 H}{\partial p_i \partial x_i} = 0. \quad (15)$$

Since K has no explicit time dependence, we have that

$$\frac{dK}{dt} = \sum_i \dot{x}_i \frac{\partial K}{\partial x_i} = \sum_i \frac{\partial H}{\partial p_i} \frac{\partial K}{\partial x_i}. \quad (16)$$

Using this result in Eq. (15) we find a differential equation for K ,

$$\frac{dK}{dt} = - \left[\sum_i \frac{\partial^2 H}{\partial p_i \partial x_i} + \frac{1}{2} \sum_{i,j} \frac{\partial^2 S}{\partial x_i \partial x_j} \frac{\partial^2 H}{\partial p_i \partial p_j} \right] K. \quad (17)$$

To find $K(\mathbf{x})$ from this equation, we need to know the Hessian, $Z_{ij} \equiv \partial^2 S / \partial x_i \partial x_j$. It is important to realise that we are only interested in quantities which vary along

the trajectories in phase space found by solving Hamilton's equations, so the momenta are now functions of \mathbf{x} . Differentiating $H = 0$ twice with respect to \mathbf{x} taking care to distinguish between the cases where \mathbf{x} and \mathbf{p} are independent and when they are not, one finds the following equation for the matrix Z [20, 21, 29, 30]:

$$\dot{Z}_{ij} + \sum_{k,l} \frac{\partial^2 H}{\partial p_k \partial p_l} Z_{ik} Z_{jl} + \sum_l \frac{\partial^2 H}{\partial x_i \partial p_l} Z_{jl} + \sum_l \frac{\partial^2 H}{\partial x_j \partial p_l} Z_{il} + \frac{\partial^2 H}{\partial x_i \partial x_j} = 0. \quad (18)$$

This can be written as a matrix equation

$$\dot{Z} + Z H_{\mathbf{pp}} Z + H_{\mathbf{xp}} Z + Z H_{\mathbf{px}} + H_{\mathbf{xx}} = 0, \quad (19)$$

where $H_{\mathbf{ab}} = \partial_{\mathbf{a}} \partial_{\mathbf{b}} H$, with \mathbf{a} and \mathbf{b} being \mathbf{x} or \mathbf{p} .

2.4. Gaussian asymptote

To actually solve Hamilton's equations and thus calculate $S(\mathbf{x})$ and $K(\mathbf{x})$ we need to know boundary conditions. The fact that the stationary distribution is approximately Gaussian near the fixed point, \mathbf{x}^* , can provide these [14, 19]. From Equation (17) we see that the differential equation which determines K is linear, and so although the dependence on \mathbf{x} can be found from this equation, K will only be known up to an overall constant.

From Eq. (9) we see that $S(\mathbf{x}^*) = 0$ and so the WKB approximation (5) can equivalently be written as

$$\pi(\mathbf{x}) = K(\mathbf{x}) \exp(-N[S(\mathbf{x}) - S(\mathbf{x}^*)]) \left[1 + \mathcal{O}\left(\frac{1}{N}\right)\right]. \quad (20)$$

Expanding about the fixed point for large N to Gaussian order gives

$$\pi(\mathbf{x}) = K(\mathbf{x}^*) \exp\left[-\frac{N}{2}(\mathbf{x} - \mathbf{x}^*)^T Z^*(\mathbf{x} - \mathbf{x}^*)\right] \left[1 + \mathcal{O}\left(\frac{1}{N}\right)\right]. \quad (21)$$

The value of $K(\mathbf{x}^*)$ now follows by asking that the probability distribution is normalised. Carrying out the Gaussian integrals one finds that

$$K(\mathbf{x}^*) = \frac{2N}{\pi} \det Z^*. \quad (22)$$

The matrix Z at the fixed point, $\mathbf{x} = \mathbf{x}^*$ and $\mathbf{p} = 0$, can be found from Eq. (19). Since $\dot{Z} = 0$ and

$$\begin{aligned} \left. \frac{\partial^2 H}{\partial x_i \partial p_j} \right|_{\mathbf{p}=0, \mathbf{x}=\mathbf{x}^*} &= \sum_{\mathbf{r}} r_j \left. \frac{\partial w_{\mathbf{r}}(\mathbf{x})}{\partial x_i} \right|_{\mathbf{x}=\mathbf{x}^*} \equiv A_{ji} \\ \left. \frac{\partial^2 H}{\partial p_i \partial p_j} \right|_{\mathbf{p}=0, \mathbf{x}=\mathbf{x}^*} &= \sum_{\mathbf{r}} w_{\mathbf{r}}(\mathbf{x}^*) r_i r_j \equiv B_{ij} \\ \left. \frac{\partial^2 H}{\partial x_i \partial x_j} \right|_{\mathbf{p}=0, \mathbf{x}=\mathbf{x}^*} &= 0, \end{aligned} \quad (23)$$

we have from Eq. (19) that $Z^* B Z^* + A^T Z^* + Z^* A = 0$, or

$$B + \Xi A^T + A \Xi = 0, \quad (24)$$

where Ξ is the covariance matrix and $Z^* \equiv \Xi^{-1}$. It should be noted that the Gaussian result (21) is what would be obtained from the system-size expansion to first order [14].

3. Solution procedure

In this section we describe how we calculate the stationary probability density from the WKB approximation. As we are primarily concerned with how the form of the stationary distribution changes with N , we fix the basic parameters as: $\beta = 1.3$, $\gamma = 1/13$, $\mu = 5.5 \times 10^{-5}$ and $\eta = 10^{-6}$. These correspond to the childhood disease measles and lead to highly oscillatory stochastic dynamics [12] (see Figure 1). The effect of changing these parameters is discussed later in Section 5.

The procedure starts with the Hamiltonian for the SIR model (11). The problem is to find the fluctuational solutions of Hamilton's equations (8). As already discussed, there is a single stable fixed point at $\mathbf{x} = \mathbf{x}^*$, $\mathbf{p} = 0$, and the non-trivial fluctuational trajectories that we seek have $\mathbf{p} \neq 0$. We can use the results of Section 2.4 to derive a consistent set of initial conditions to equations (8) by noticing that $S(\mathbf{x})$ is quadratic in $\mathbf{x} - \mathbf{x}^*$ near the fixed point [19]. If we define a small perturbation, δx_j , to the fixed point, then the corresponding initial perturbations to the momentum along the required trajectory are given by

$$\delta p_i = \sum_j \left. \frac{\partial p_i}{\partial x_j} \right|_{\mathbf{p}=0, \mathbf{x}=\mathbf{x}^*} \delta x_j = \sum_j Z_{ij}^* \delta x_j = \sum_j (\Xi^{-1})_{ij} \delta x_j. \quad (25)$$

The covariance matrix Ξ may be found from Equation (24), since A and B can be found from Equation (23). An example of a fluctuational trajectory found via this procedure is shown in Figure 2. The trajectory spirals away from the fixed point and ends when it leaves the positive quadrant of the x - y plane. Some care has to be taken when specifying the initial perturbations; they must be small enough that Equation (25) applies, but not so small that numerical errors can build up along the trajectory, due to the exponential slow down near the fixed point. For the same reason there is a practical limit to the range over which trajectories, and hence $\pi(\mathbf{x})$ can be calculated with this method.

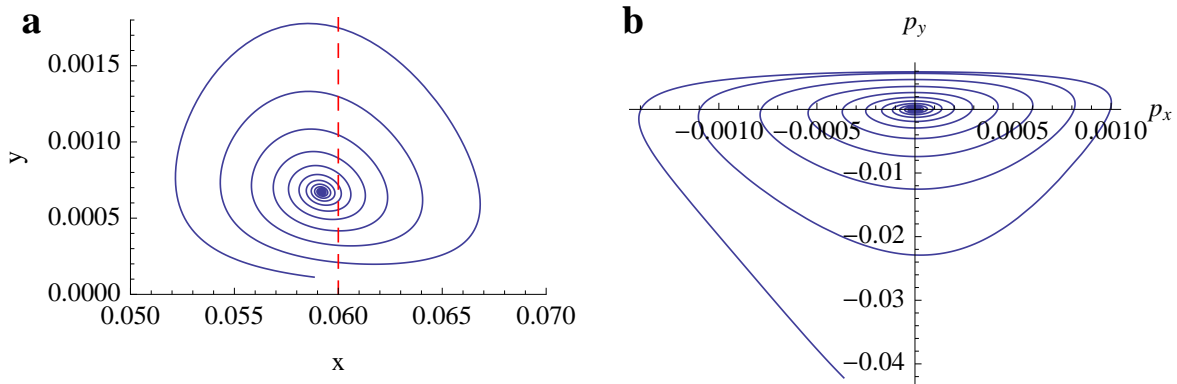


Figure 2. An example of a fluctuational trajectory emanating from the fixed point $\mathbf{x} = \mathbf{x}^*$, $\mathbf{p} = 0$, projected into the (a) x - y plane, and (b) p_x - p_y plane. The line $x = 0.06$ is marked by the red dashed line.

With the initial conditions, plus $S(\mathbf{x}^*) = 0$, $K(\mathbf{x}^*) = 1$ and $Z(\mathbf{x}^*) = \Xi^{-1}$, we can

simply numerically integrate forward the equations (8), (9), (17) and (18) simultaneously to find a trajectory as well as $S(\mathbf{x})$ and $K(\mathbf{x})$ along it. We have taken $K(\mathbf{x}^*) = 1$, rather than the value given by Equation (22) for convenience. The final results are always normalised numerically, so the choice we make here is not important. By changing the initial perturbation, δx_j , different trajectories can be found. As the functions S and K , and hence π , are only given along the trajectory, we need a method to interpolate these into a distribution over the x - y plane.

In principle one can use a shooting method to find a trajectory which goes through any given point in the phase space, and hence the probability at that point. In practice, because of the highly oscillatory nature of the solutions, it is easier to calculate cross-sections of the distribution by calculating where a large set of trajectories intersect a given line in the x - y plane. The set of trajectories are computed by using initial conditions which lie on a small circle surrounding the fixed point [19]. In this way a discretised estimate of $\pi(\mathbf{x})$ can be computed. Figure 3a shows the functions $S(\mathbf{x})$ and $K(\mathbf{x})$ evaluated where trajectories intersect the line $x = 0.06$. The probability density along this cross section is then given by $P_{\mathbf{n}} \equiv \pi(\mathbf{x}) = K(\mathbf{x}) \exp[-NS(\mathbf{x})]$, shown in Figure 3b. Simulation results are not shown in Figure 3, but are in excellent agreement with the analytic results. Although there is no reason to suppose that these functions have simple analytic forms, one could ask if a simple analytic expression can be fitted in the asymptotic regime $y \gg y^*$. We have observed that in this regime good fits to the functions shown in Figure 3a are $K(y) \sim a/(b + y)$ and $S(y) \sim (cy + d)$, for some constants a, b, c and d . There is, however, no analytic justification for these.

For comparison, Figure 3c shows π with and without the pre-factor K . Clearly for smaller systems, this contribution from the next-to-leading order terms in the expansion cannot be ignored when finding an accurate estimate of π . More generally, we wish to calculate the full marginal infective probability density, $P_m = \sum_n P_{n,m}$. This is achieved by simply calculating the probability density along a range of cross-sections and then summing over them. The resulting marginal density is then normalised such that $\sum_m P_m = 1$, as it ought to be.

4. Outbreak distributions

In this section we present results which show how the outbreak distribution changes with population size, N . The mean incidence varies linearly according to $m^* = Ny^*$. As the mean is decreased, three dynamical regimes can be distinguished, classified by population size, which coincide roughly with type I, II and III dynamics originally suggested by Bartlett [4]. Type I systems are large; they are those in which fade-out does not occur on a practical timescale, thus the disease is endemic. For the parameters used in this paper the lower limit on this class is $N \sim 8 \times 10^5$. For type II, medium size systems, fade-out can occur but reintroduction is swift and the dynamics are still mainly endemic. Type III are small systems which cannot support endemic infection levels. In these systems outbreaks are triggered by immigration events, which then rapidly lead to

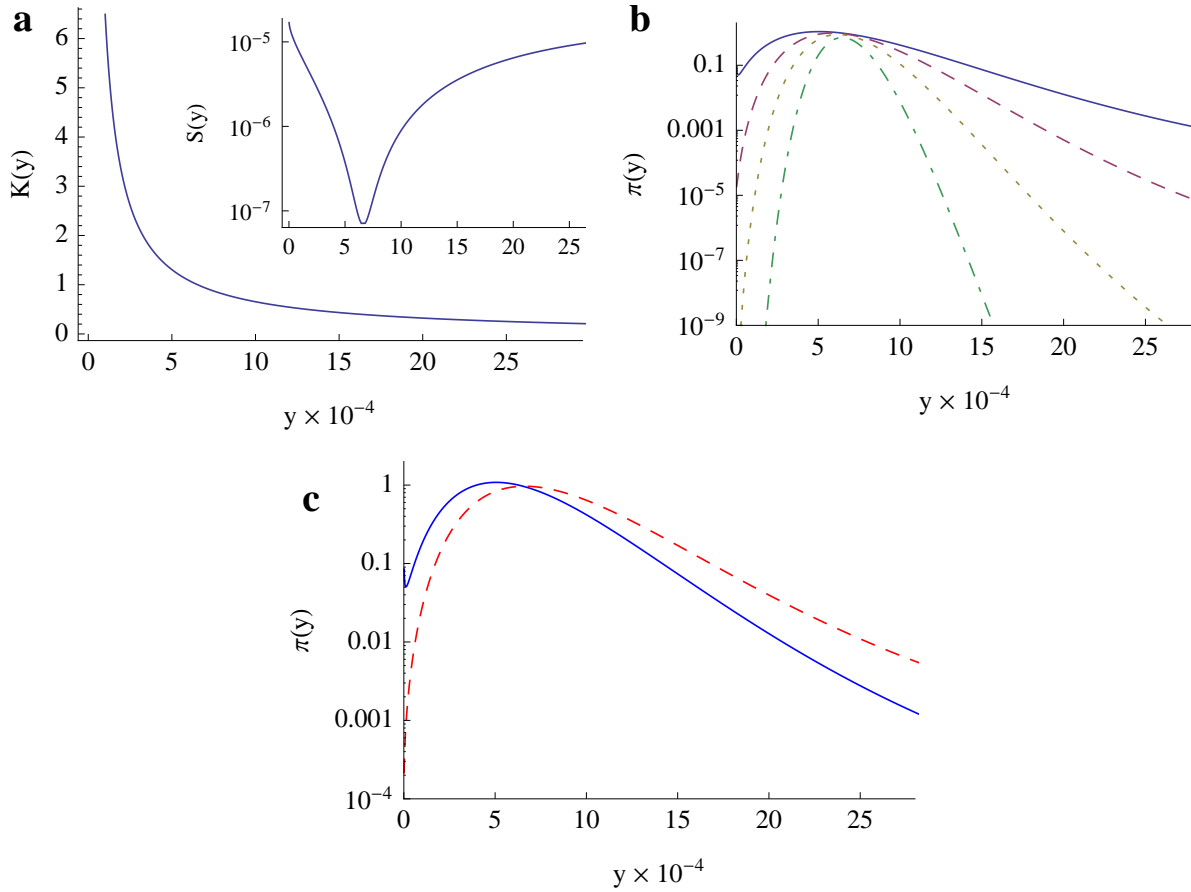


Figure 3. (a) The functions K and S evaluated along the line $x = 0.06$ (see Figure 2a). (b) The probability density, $\pi(y) = K(y) \exp[-NS(y)]$, evaluated for $N = 5 \times 10^6$, 2×10^6 , 10^6 and 5×10^5 (solid, dashed, dotted and dot-dashed lines respectively). (c) Comparison of $\pi(y)$ calculated with (solid blue curve) and without (red dashed curve) the pre-factor $K(y)$; $N = 5 \times 10^{-5}$. The $\pi(y)$ are left un-normalised in this figure for clarity.

fade-out. Three time series representative of these regimes are shown in Figure 1. The marginal infective distribution, P_m , then describes the distribution of outbreak sizes. We will primarily be interested in how this changes with N , for type I and II systems.

4.1. Type I systems

Figure 4 shows the outbreak distributions, calculated via the WKB approximation and compared to simulations, for three large populations in which the probability of extinction is negligible. For all three the agreement between the result from the WKB calculation and the stochastic simulations is excellent. As N is decreased so is the mean (indicated by the arrows) and the distribution becomes more asymmetric as the fade-out boundary starts to exert a bigger influence on the dynamics. These distributions are also compared with the Gaussian result given by Equation (21) (red dashed lines). It can be seen that this overestimates the probability of the number of infected becoming

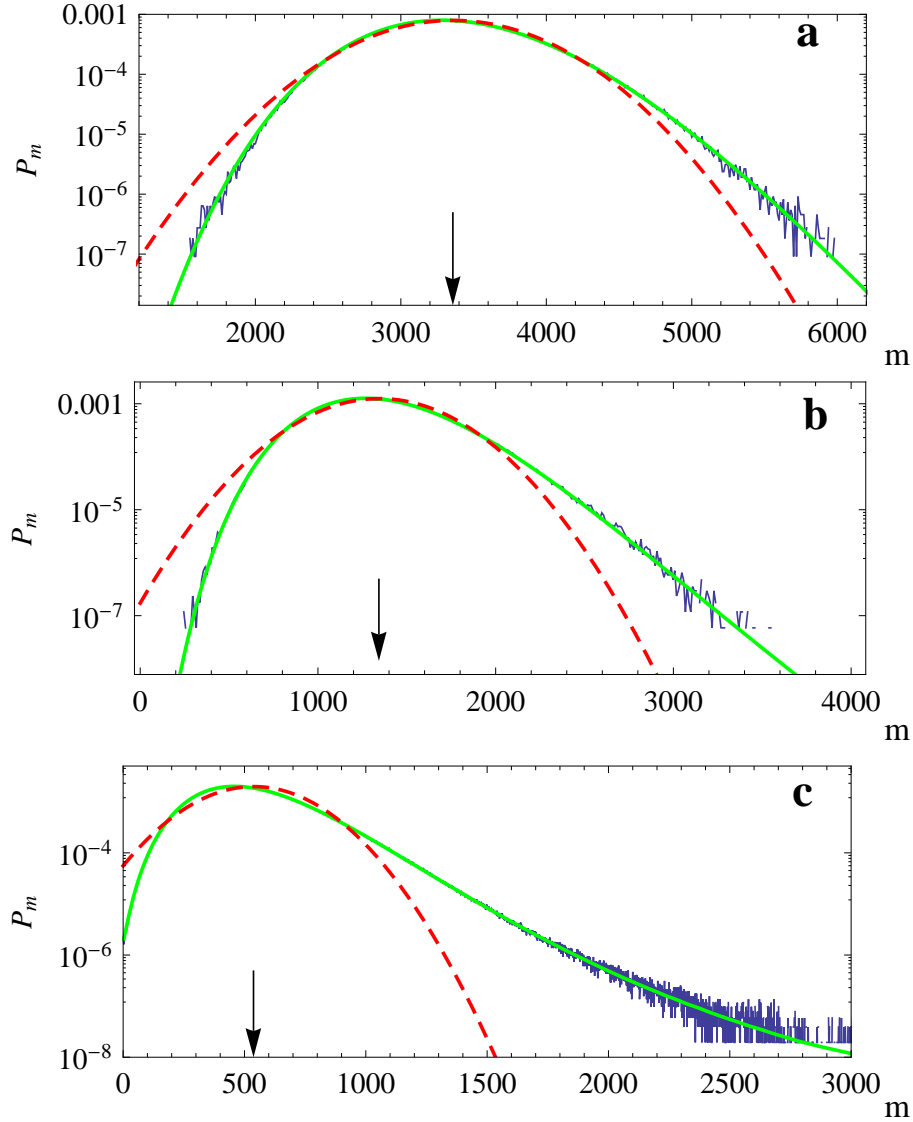


Figure 4. Marginal infective probability densities for larger systems. Population sizes are (a) $N = 5 \times 10^6$, (b) $N = 2 \times 10^6$ and (c) $N = 8 \times 10^5$. Solid green lines are from WKB calculations, the red dashed lines show the Gaussian approximation, and the noisy blue lines are from stochastic simulations. The arrows mark the mean values.

small, while severely underestimating the probability of very large outbreaks. The WKB approximation remains valid down to $m \sim \mathcal{O}(1)$, but diverges close to the boundary as expected [27].

The fluctuations when m is small (usually after an outbreak) determine the size of the next outbreak. If a fluctuation takes the system closer to the boundary then the probability that the next outbreak will be large is greatly increased. This is because if m remains small, a large pool of susceptibles can build up. This effect is illustrated in Figure 5, which shows how the outbreak distribution changes with initial conditions close to and further away from the boundary. Figure 5b also shows that larger outbreaks have a naturally longer period. This feature accounts for the observed changes in the

power spectrum of fluctuations in smaller systems, which are systematically shifted to lower frequencies, as compared with the $N \rightarrow \infty$ predictions [11, 31].

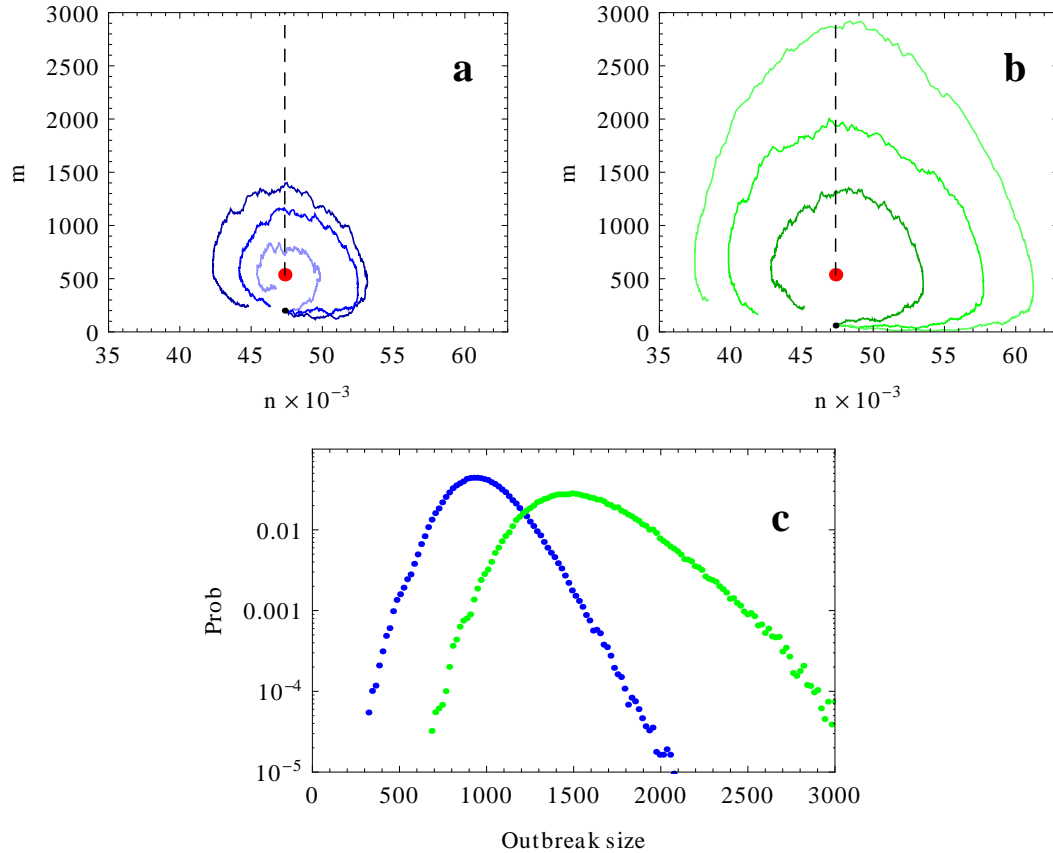


Figure 5. The effect of the boundary on outbreak size. Parts (a) and (b) show realisations of the stochastic model started from $\mathbf{x}_0 = (Nx^*, 200)$ and $\mathbf{x}_0 = (Nx^*, 60)$ respectively, with $N = 8 \times 10^5$. Dynamics then proceed in an anti-clockwise fashion. The realisations that fluctuate nearer to the boundary tend to move further along the n -axis leading to larger outbreaks. Part (c) shows the outbreak distributions—defined as when the realisation first crosses the dashed lines—for the two initial conditions.

4.2. Type II and III systems

For the population sizes considered in the previous section, fade-out for any length of time is very rare, but even so, the proximity of the boundary and the potentially low levels of infectious individuals can have a large effect on the dynamics. We now study two smaller systems, and show how fade-outs of the disease affect the dynamics, and where the WKB approximation starts to break down. The outbreak distribution for $N = 5 \times 10^5$ is shown in Figure 6. The inset shows the asymptote of the distribution near the fixed point and there is clearly a non-zero probability of being on the boundary. In spite of this, the WKB result is still in very good agreement with the simulation result.

For smaller populations this agreement starts to break down. This can be seen in Figure 7 which shows the outbreak distribution for $N = 3 \times 10^5$. Although the WKB

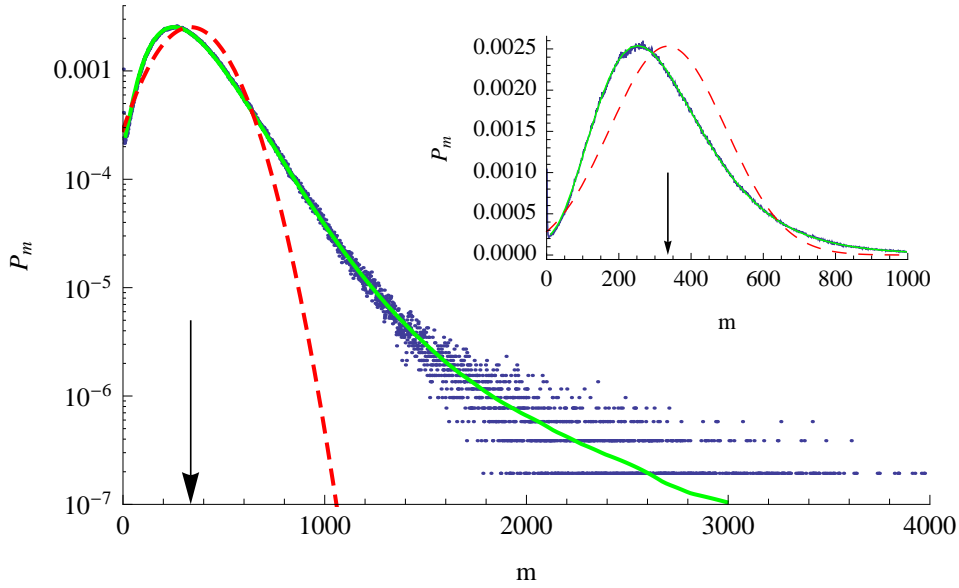


Figure 6. Outbreak distribution for $N = 5 \times 10^5$. Solid green line is the WKB result, red dashed line is the Gaussian approximation and the blue dots are from simulations. Inset shows the asymptote around the fixed point; the probability of fade out is now clearly non-zero. The arrow marks the mean value.

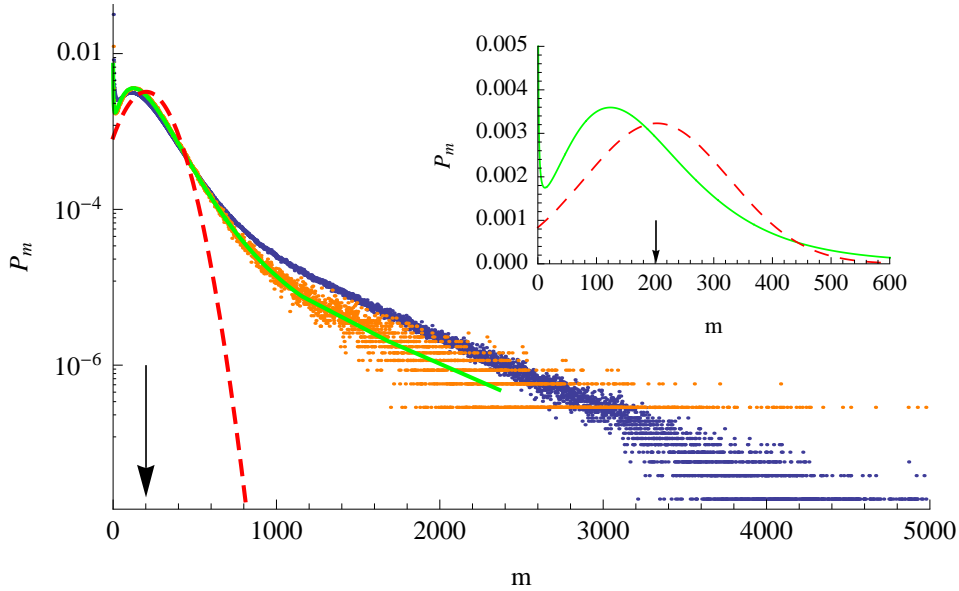


Figure 7. Comparison of the full probability density (blue dots) and a single cross-section through the middle of the distribution (orange dots) for $N = 3 \times 10^5$. Again the green line, red dashed line and arrow show the WKB, Gaussian and mean result respectively. At this system size the WKB approximation cannot capture the full distribution, but it does approximate well the central cross-section. The inset shows the asymptote of the cross-section about the fixed point, which shows that there is significant build up of probability on the boundary.

method provides a good approximation to a cross-section through the middle of the distribution, it cannot correctly predict the tails of the full marginal distribution, which is now strongly bimodal. This is because a large contribution to these tails comes from outbreaks which are started *on* the boundary. Thus to estimate the tails by the WKB method one would need to find trajectories which start on the boundary. This might in principle be possible—if we could generate initial conditions.

For the larger system in Figure 6 there are also outbreaks which originate from the boundary but their contribution to the overall marginal density can still be calculated from trajectories started at the fixed point. This is because the system never moves too far along the boundary before an immigration event starts a new outbreak. For $N = 3 \times 10^5$ this is no longer the case as the system can potentially move far along the boundary into a region where it is no longer possible to find trajectories via the normal methods.

The dynamics of even smaller systems are dominated by fade-out at the boundary as shown in Figure 1c. In effect, almost all outbreaks are started from the boundary and are of such a large size that they almost always lead to fade-out. Susceptibles then build up through births until an immigration event triggers another outbreak. Therefore the timing of an outbreak is a random process, with the average outbreak size related to the time since the last outbreak (and hence the size of the pool of susceptibles).

5. Discussion

In this article we have analytically investigated the outbreak distribution of a stochastic SIR model with immigration. This is carried out by calculating both the leading order exponential, and next-to-leading order pre-factor in a WKB expansion of the master equation. The agreement of these analytic results with stochastic simulations is excellent. Although we cannot calculate the full marginal distribution for very small systems, where fade-out starts to dominate the dynamics, the WKB calculation is still remarkably accurate when calculating cross-sections of the bulk of the distribution.

The ability of smaller systems—those where the number of infected individuals can become small—to generate relatively larger outbreaks is clear in simulations and is important for analysing real time series. Up until now this effect has received little theoretical attention. In the limit $N \rightarrow \infty$ the distribution is Gaussian, a fact which is the basis of the system-size expansion [14]. But as N is made smaller, the distribution becomes highly non-Gaussian as the fade-out boundary approaches the macroscopic fixed point. Although the asymptote of the distribution remains somewhat Gaussian in shape, the fat tails of the distribution indicate the presence of large outbreaks (relative to the mean infectious level). These tails cannot be captured within the Gaussian approximation. Clearly these results are relevant to a number of models of this type; the SIR model is just one example of a generalised predator-prey model [32].

For smaller systems, the next-to-leading order term makes a significant contribution to the calculation of the probability density (see Figure 3c). Most studies of two-

dimensional systems up until now have obtained approximate results for fixation times in various limits where Hamilton's equations can be simplified and solved analytically [23–25]. Although we have not been concerned with fixation in this paper, our work has implications for calculating the mean time to extinction using these methods. By definition, if a system is small enough that extinction can happen on an observable time scale, then there will be large corrections to the first-order result from the next-to-leading order terms in the WKB expansion.

Throughout this paper we have fixed the basic parameters of the system corresponding to a disease like measles, which has highly oscillatory stochastic dynamics and is typical of many childhood diseases [12]. Our aim has been to illustrate the method which we use, giving enough detail so that others can use it for other parameter values or other models. The choice of parameters was therefore not chosen on theoretical grounds; in principle any set which showed similar effects could have been used. While this is true for the standard SIR parameters β , γ and μ , it is worthwhile stressing the role of the parameter η , which is less standard. This parameter governs the rate of new infections brought back by susceptibles visiting outside populations. In the natural dynamics this can be thought of as a damping term because it is constantly depleting the pool of susceptibles, thus the larger η is set, the smaller and less coherent the size of the stochastic oscillations for a given population size [8]. Such a system will show the same overall changes in the form of the outbreak distribution, but the onset will be later at comparatively smaller system sizes.

One of the factors we have ignored in this paper is the influence of seasonal forcing on the epidemic dynamics. This is known to be very important for childhood diseases and arises due to the aggregation of children in schools during term time. To include this, the transmission term β is made a function of time with a period of one year. The forcing then induces a limit cycle in the macroscopic dynamics with stochastic amplification of the transients about this [11]. One of the most interesting and biologically relevant features of these forced models is the existence of multiple stable attractors; for example the coexistence of 1 and 2-year cycles [33,34]. The WKB methods outlined in this article should be suitable for analysing the rate of switching between these attractors and for showing how this changes with system size [35]. A quantitative stochastic theory of this phenomenon would be a valuable addition to our overall understanding of these epidemic systems.

Acknowledgments

A.J.B. acknowledges support from the EPSRC and the Australian Research Council's Discovery Projects funding scheme (project number DP110102893).

References

- [1] Bauch C T 2008 The role of mathematical models in explaining recurrent outbreaks of infectious childhood diseases *Mathematical Epidemiology* ed Brauer F, van den Driessche P and Wu J

- (Springer) pp 297–319
- [2] Metcalf C J E, Bjornstad O N, Grenfell B T and Andreasen V 2009 *Proc. R. Soc. B* **276** 4111–4118
 - [3] Rozhnova G and Nunes A 2010 *Phys. Rev. E* **82** 041906
 - [4] Bartlett M S 1957 *J. R. Stat. Soc. Ser. A* **120** 48–70
 - [5] Rohani P, Earn D J D and Grenfell B T 1999 *Science* **286** 968–971
 - [6] Bauch C T and Earn D J D 2003 *Fields Inst. Comm.* **36** 33–44
 - [7] Keeling M J and Rohani P 2007 *Modelling Infectious Diseases in Humans and Animals* (Princeton: Princeton University Press)
 - [8] Alonso D, McKane A J and Pascual M 2007 *J. R. Soc. Interface* **4** 575–582
 - [9] Bauch C T and Earn D J D 2003 *Proc. R. Soc. Lond. B* **270** 1573–1578
 - [10] Black A J and McKane A J 2010 *J. R. Soc. Interface* **7** 1219–27
 - [11] Black A J and McKane A J 2010 *J. Theor. Biol.* **267** 85–94
 - [12] Anderson R M and May R M 1991 *Infectious Diseases of Humans: Dynamics and Control* (Oxford: Oxford University Press)
 - [13] Grenfell B and Bolker B 1998 *Ecol. Lett.* **1** 63–70
 - [14] van Kampen N G 1992 *Stochastic Processes in Physics and Chemistry* (Amsterdam: Elsevier)
 - [15] Kramers H A 1940 *Physica* **7** 284–304
 - [16] Goldstein H, Poole C P and Safko J L 2001 *Classical Mechanics* 3rd ed (San Francisco: Addison-Wesley)
 - [17] Kubo R, Matso K and Kitahara K 1973 *J. Stat. Phys.* **9** 51–96
 - [18] Gang H 1987 *Phys. Rev. A* **36** 5782–5790
 - [19] Dykman M I, Mori E, Ross J and Hunt P M 1994 *J. Chem. Phys.* **100** 5735–5750
 - [20] Stein D L and Maier R S 1997 *SIAM J. Appl. Math.* **57** 752–790
 - [21] Roma D, Oflanagan R, Ruckenstein A, Sengupta A and Mukhopadhyay R 2005 *Phys. Rev. E* **71** 011902
 - [22] Schwartz I B, Billings L, Dykman M and Landsman A 2009 *J. Stat. Mech.* P01005
 - [23] Meerson B and Sasorov P 2009 *Phys. Rev. E* **80** 41130
 - [24] Kamenev A and Meerson B 2008 *Phys. Rev. E* **77** 61107
 - [25] Khasin M, Dykman M I and Meerson B 2010 *Phys. Rev. E* **81** 051925
 - [26] Escudero C and Kamenev A 2009 *Phys. Rev. E* **79** 041149
 - [27] Assaf M and Meerson B 2010 *Phys. Rev. E* **81** 021116
 - [28] Risken H 1989 *The Fokker-Planck Equation* 2nd ed (Berlin: Springer)
 - [29] Maier R and Stein D 1993 *Phys. Rev. Lett.* **71** 1783–1786
 - [30] Bressloff P 2010 *Phys. Rev. E* **82** 051903
 - [31] Simoes M, da Gama M M and Nunes A 2008 *J. R. Soc. Interface* **5** 555–566
 - [32] Brauer F and Castillo-Chavez C 2001 *Mathematical Models in Population Biology and Epidemiology* (New York: Springer)
 - [33] Earn D, Rohani P, Bolker B and Grenfell B 2000 *Science* **287** 667–670
 - [34] Schwartz I B 1985 *J. Math. Biol.* **21** 347–361
 - [35] Smelyanskiy V N, Dykman M I and Maier R S 1997 *Phys. Rev. E* **55** 2369–2391